PROTOCOL EXCHANGE | COMMUNITY CONTRIBUTED

Synthesis and screening of

Tom Ellis

Jim Collins Group (BU)

Introduction

Parallel-synthesis of biomolecular components offers a rapid path to constructing multiple gene networks that behave as predicted by *in silico* modeling. Here we describe the synthesis and screening steps to produce a component library, using the example of repressible promoters for *S. cerevisiae*. Regulatory promoters such as these are crucial components in genetic networks and quantified variability in input and output can be used to design diverse networks. The parallel synthesis approach described here is not limited to promoters or *S. cerevisiae* and should be broadly applicable to all biomolecular components with protocol adaptations.

Three main methods exist for producing component libraries; (i) combinatorial shuffling (Cox et al., 2007; Gertz et al., 2009; Ligr et al., 2006; Murphy et al., 2007), (ii) multi-round mutation/selection of an existing component (Alper et al., 2005; Alper et al., 2006), and (iii) diversity-inherent component synthesis (Jensen and Hammer, 1998a, b). We describe here the third method, which we recommend as it is rapid and compatible with rational design of new components.

Subject terms: Nucleic acid based molecular biology

Keywords: components parts libraries screening modeling synthetic biology synthesis characterization regulation promoter yeast

Procedure

1. Design of biomolecular component

For diversity-inherent component synthesis it is recommended to base the component design on an existing well-characterized component, where essential motifs are understood and consensus sequences are available. Extra motifs (e.g. for regulation) can be introduced to this component, and if designed correctly, should behave as intended. It is also possible with this synthesis technique to create an entirely new hybrid component by linking together several characterized motifs in a rational design.

For the design, the core sequences of essential motifs are specified and the base sequences between these are unspecified (i.e. defined as ‘N’). It is recommended that the precise nucleotide distance between motifs is maintained; otherwise their function can be lost. More ‘N’ sequence in the design (including some within motifs) ensures more variation in the final component library but will require more screening as a higher percentage of components risk being non-functional.

Example 1: *E. coli* LacXYZ promoter
-35 -10 +1 *Olac*
cacccca **ggctttaca** ctttatgcttccggct **cgtatgtt** gtgtg **gAattgtgagcggataacaat** ttcac

nnnnnnn **ggctttaca** nnnnnnnnnnnnnnn **cgtatgtt** nnnnn **gAattgtgagcggataacaat** nnnnn

Example 2: *H. Sapiens* Myc translation start (mRNA)
5'UTR Kozak Seq Myc CDS
gacgctggatttttttcgggtagtggaaaaccagcagcct **cccgcgacgATGC** CCCTCAACGTTAGCTTC

gacgctgnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn **nnnnncaccATGC** CCCTCAACGTTAGCTTC

A fundamental concern when designing the component is the presence of restriction enzyme sites. The component should be designed to be flanked by two reliable restriction enzyme sites that can be cut together at high efficiency (a double digest) to give good sticky ends for efficient ligation. Between these restriction enzyme sites, the ideal design will contain none of the restriction enzyme sites that are used to exchange modules of the characterization vector (**step 3**).

2. *In vitro* synthesis of component library
Several methods exist for physically putting together the DNA for the component; (i) custom synthesis order from a gene synthesis company, (ii) *Taq* polymerase extension from one long primer (Solem and Jensen, 2002), and (iii) *Klenow* pol extension of two annealed long oligonucleotides (Jensen and Hammer, 1998b). Described here is the third method, which allows for ~200 bp of diversified sequence from two affordable oligonucleotides. A schematic is shown in **Figure 1**.

a) Design oligonucleotides, defining each unspecified nucleotide as 'N'. To allow annealing of oligos, a region of 10 to 20 conserved base pairs is required to fall about halfway in the component design and each oligo will end with part or all of this sequence. This can usually be made to coincide with a fixed motif such as a binding site. It is also essential to add at least three nucleotides 5' to any restriction enzyme sites that begin each oligo.

b) Order oligonucleotides – above 100 nt length it is essential to get PAGE-purification and usually required to order a higher yield. Upon receipt, dilute oligos in sterile water to 10 µM.

c) In a PCR tube mix 2 µl of each oligo, 10 µl of NEBuffer 2 (NEB, Beverly MA), 10 µl of 10x BSA and 75 µl of sterile water. Float the tube in a bath of fast-boiling water for 5 minutes and then leave the water bath and tube to cool down to slowly room temperature (~2 hours).

d) Add of 1 µl DNA Polymerase I, Large (Klenow) Fragment (NEB, Beverly MA) and 1 µl 0.25 mM dNTP mix (NEB, Beverly MA) and incubate for 1 hour at 37°C. Inactivate the enzyme with 10 minutes at 75°C and cool slowly to room temperature.

e) Purify the correct sized double-stranded DNA band using agarose gel electrophoresis and QIAquick Gel Purification kit (Qiagen, Valencia, CA).

3. Design of characterization vector
To screen and characterize the component in the host system requires a vector to introduce and maintain

the component in the system as well as any other genetic material needed to assess the component. Having commonly-used unique restriction enzyme sites between features of the vector allows for modularity.

For the example of a repressible promoter library, the vector requires (i) a reporter gene coding sequence downstream of the library promoter cloning site, (ii) a null sequence at the library promoter cloning site, (iii) the repressor gene coding sequence under the control of a constitutive promoter, (iv) the usual genes and sequences for selecting transformants and vector propagation in bacteria.

For working in *S.cerevisiae* an auxotrophic marker and site for integration into the yeast genome are required (Amberg et al., 2005) and it is also advisable to have terminator sequences following the repressor and reporter genes. To ensure clearing of histones from the library promoter region, the *GAL1* upstream activator sequence (UAS) can be included upstream and yeast experiments performed in galactose. An example vector is shown in **Figure 2**.

4. Introduction of component library into host system

a) Perform a double-digest reaction on the vector and the *in vitro* promoter library and clean-up the products using QIAquick Nucleotide Removal Kit (Qiagen, Valencia, CA) for the promoter library and performing agarose gel electrophoresis, and QIAquick Gel Purification (Qiagen, Valencia, CA) for the vector piece.

b) Ligate vector and insert using T4 DNA Ligase (NEB, Beverly MA) according to the manufacturer's guidelines.

c) Large-scale transform *E. coli* cells with unpurified ligation products. Electrocompetent or chemically competent cells are both usable, but the highest efficiency possible is desired to get a good library. High-efficiency chemically-competent *DH5α E. coli* (NEB, Beverly MA) are recommended, using 4 reaction aliquots per library and plating out transformation products over 16 LB+Amp agar plates (4 plates per transformation). Follow transformation protocol accompanying the competent cells, and also include a negative control, plating 1/4 of this transformation on a single plate. Leave for around 12 hours at 37°C until colonies are clearly visible but still too small to be confluent.

d) If transformation is successful (i.e. number of colonies on each plate is significantly greater than that on the control plate) then these plates are the bacteria library and if intended for use in *E. coli*, then the clones are ready to be screened.

e) For work in other systems collect all of the colonies by scraping off each plate into 1ml LB using a sterile plate spreader. Pool the colonies in a single tube, spin down the cells, extract all plasmid into sterile water using a QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA) and determine the yield and DNA concentration. This is the plasmid library.

f) The pooled plasmid library can be used to transform the host system. Yeast transformation is described in detail elsewhere(Amberg et al., 2005). For libraries, a restriction enzyme digestion at the integration site is performed for 1 µg of pooled plasmid. Clean this digest reaction, eluting into water using a QIAquick purification (Qiagen, Valencia, CA) and use this linearized DNA to transform yeast following a standard LiAc protocol but scaled up to by 10 times and plating out on two 250 mm x 250 mm

auxotrophic marker selective plates. To check transformation efficiency, a single negative control transformation with pure water can be performed and plated onto a single 70 mm diameter plate. Incubate plates for 48 hours at 30°C.

g) If transformation is successful (i.e. thousands of colonies on the two library plates and almost none on the negative control), then these plates form the yeast library and clones are ready to be screened.

5. Screening of component library

Screening the component library requires taking individual colonies from the library plates, growing these up and assaying them for intended function. By using 96-well format methods, a large number of colonies can be screened very quickly.

For the example of yeast repressible promoters the procedure is as follows:
a) In each well of two 96-well plates, inoculate 300 µl of ON-state growth media with a single colony from the library and incubate for 22 hours at 30°C. For TetR-regulated promoters with a *GAL1* UAS, ON-state growth media is YPD plus 2% galactose with a saturating concentration of TetR inhibitor (250 ng/ml ATc). YPD (also known as YEPD) is 1% yeast extract, 2% peptone in purified water (autoclaved).

b) Estimate output of ON-state promoter by parallel measurement of reporter gene output. For yEGFP, cell fluorescence is measured at 450 nm using a fluorescence-capable Plate Reader (e.g. SpectraFluor Plate Reader, Tecan, Durham NC).

c) In each well of a new 96-well plate, inoculate 300 µl of OFF-state growth media with every well from (b) that showed measurable fluorescence (typically 20-40% of wells). Then incubate for 22 hours at 30°C. For TetR-regulated promoters with a *GAL1* UAS, OFF-state growth media is YPD plus 2% galactose.

d) Estimate output of OFF-state promoter by parallel measurement of reporter gene output as in (b). Candidates for a repressible promoter library are those in (d) that showed a decrease in output compared to (b) (typically 60-80% of wells).

e) In each new well of a new 96-well plate, inoculate 300 µl of basic growth media with every well from (d) that was a candidate gene, and then incubate for 22 hours at 30°C. This should be about 25 to 50 wells. Basic growth media is YPD plus 2% glucose (Amberg et al., 2005). If any wells show measurable fluorescence in basic media, discard these.

f) Perform integration-check PCR (Amberg et al., 2005) from cells taken from each remaining well in (e) to screen for single-integrations at the intended genomic locus. For colony PCR of yeast, place 2 µl of cells in 50 µl of 50 U/mL lyticase, heat to 37°C for 30 min, then 95°C for 10 min, then cool to 4°C. Use 3 µl of this as template in a 30 µl PCR with 40 cycles. Discard any wells that show up as double-integrated in the PCR screen.

g) Using the lysed cells produced in (f), PCR amplify the inserted library-made DNA from each remaining well and get this sequenced. Discard any duplicates identified from sequencing. The remaining wells are all candidates for the final library.

h) Choose your final library (e.g. 20 clones) from the candidates, selecting to have as diverse as possible range of output levels in the ON and OFF state. If the data from steps (b) and (d) are low resolution, it may

be worthwhile repeating these with a higher-resolution technique (such as flow cytometry) before selecting the final candidates.

6. Characterization of component library

Characterizing the selected component library is no different to characterizing a regular component, just scaled-up and done in parallel. A library designed to have diversity in one or two properties should be assayed carefully for these, with biological repeats to ensure quality data. Theoretically, other properties should be conserved throughout the library, so only one member needs to be characterized for these. However, it may be desirable to assay a few library members together to ensure this is the case.

For the example of yeast repressible promoters the procedure is as follows:
a) In 12 ml culture tubes, inoculate 3 ml of OFF-state growth media and 3 ml of ON-state growth media for each member of the library, inoculating from cells growing in basic growth media. Incubate for 20 hours at 30°C with orbital shaking, to an approximate optical density at 600 nm (OD_{600}) of 1.00.

b) Measure the ON- and OFF-state expression for each promoter with a high-resolution assay. For fluorescent reporters, flow cytometry is ideal.

c) Repeat steps (a) and (b) to get biological repeat data. For each promoter in the library, the mean ON-state expression is the S_{max} and the mean OFF-state expression is S_{min}.

d) Repeat steps (a) to (c) for a control construct where the library-generated promoter is replaced by the same promoter sequence directing expression of the repressor protein. The output value here is essential as it gives a measure of how much repressor is present in the system in the ON and OFF states.

e) Characterize other properties of the component as appropriate, either just for a single example member or for as many as necessary. Characteristics desirable for modeling work with repressible promoters include induction curves (how promoter output changes as repressor concentration is titrated by a chemical inhibitor) and temporal activation curves (how promoter output changes over time after ON state is induced by the change of media).

Timing

Design and synthesis of DNA:
2 weeks (dependent on speed of outsourced DNA or oligo synthesis)

Library synthesis, screening and characterization:
1 to 2 weeks for bacterial component library
2 to 3 weeks for yeast component library
(Note: assembly of a new characterization vector can take several weeks)

Critical Steps

Step 1. Component design: the distance between conserved motifs should be maintained.

Step 3. Vector design: the vector **must** have a null sequence where the component DNA is going to be ligated in place of. This avoids false-positives occurring during screening.

Step 4. Ligation and transformation: to get as large a library as possible, high-efficiency ligation and transformation are required.

Troubleshooting

1. Screen yields no functional components – repeat **steps 4 and 5** with a characterized control component instead of library DNA, to check that the screen and vector work as intended.

2. Low transformation efficiency in *E. coli* – repeat library/vector ligation but only use well-cut clean DNA that has not been exposed to UV during gel-purification (UV exposure of DNA severely decreases transformation efficiency).

3. Failing integration-check colony PCR of yeast genome integrations – harvest genomic DNA from yeast clones (Amberg et al., 2005) and use this for PCR instead. Touchdown PCR is recommended.

References

Alper, H., Fischer, C., Nevoigt, E., and Stephanopoulos, G. (2005). Tuning genetic control through promoter engineering. Proceedings of the National Academy of Sciences of the United States of America 102, 12678-12683.

Alper, H., Moxley, J., Nevoigt, E., Fink, G.R., and Stephanopoulos, G. (2006). Engineering yeast transcription machinery for improved ethanol tolerance and production. Science (New York, N.Y) 314, 1565-1568.

Amberg, D.C., Burke, D., and Strathern, J.N. (2005). Methods in yeast genetics : a Cold Spring Harbor Laboratory course manual, 2005 ed. edn (Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press).

Cox, R.S., 3rd, Surette, M.G., and Elowitz, M.B. (2007). Programming gene expression with combinatorial promoters. Molecular systems biology 3, 145.

Gertz, J., Siggia, E.D., and Cohen, B.A. (2009). Analysis of combinatorial cis-regulation in synthetic and genomic promoters. Nature 457, 215-218.

Jensen, P.R., and Hammer, K. (1998a). Artificial promoters for metabolic optimization. Biotechnology and bioengineering 58, 191-195.

Jensen, P.R., and Hammer, K. (1998b). The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. Applied and environmental microbiology 64, 82-87.

Ligr, M., Siddharthan, R., Cross, F.R., and Siggia, E.D. (2006). Gene expression from random libraries of yeast promoters. Genetics 172, 2113-2122.

Murphy, K.F., Balazsi, G., and Collins, J.J. (2007). Combinatorial promoter design for engineering noisy gene expression. Proceedings of the National Academy of Sciences of the United States of America 104, 12726-12731.

Solem, C., and Jensen, P.R. (2002). Modulation of gene expression made easy. Applied and

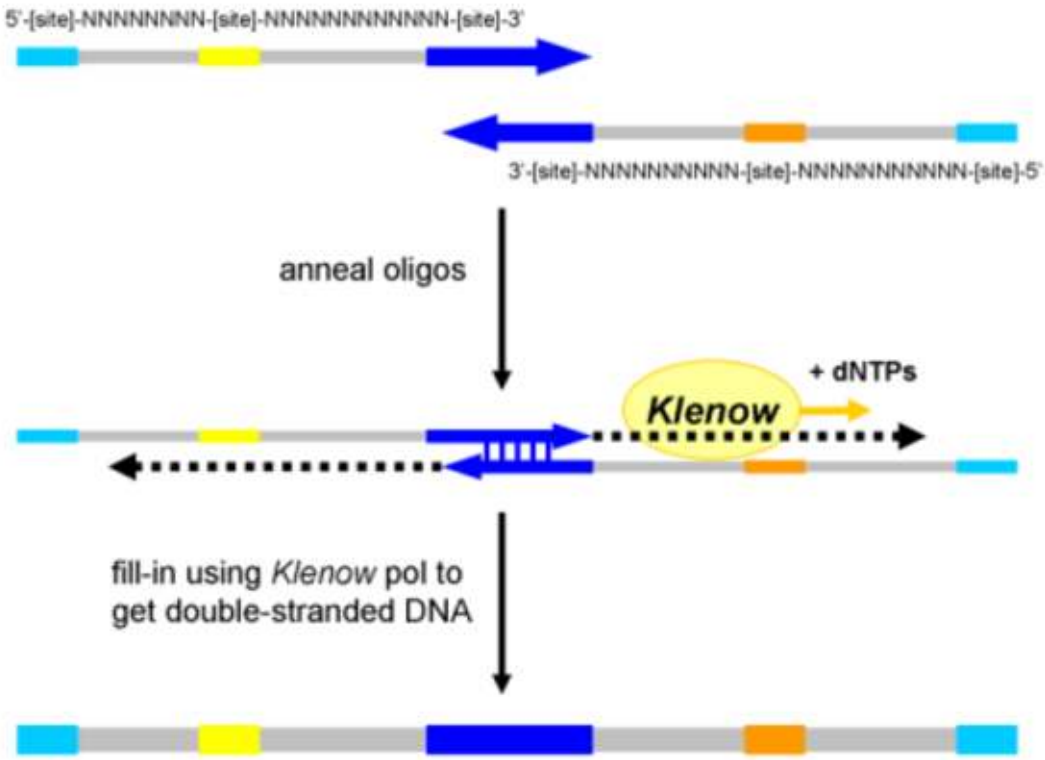
environmental microbiology 68, 2397-2403.

Acknowledgements

We thank Peter Jensen, Kevin Murphy and Ellen O'Shaughnessy for useful discussions.

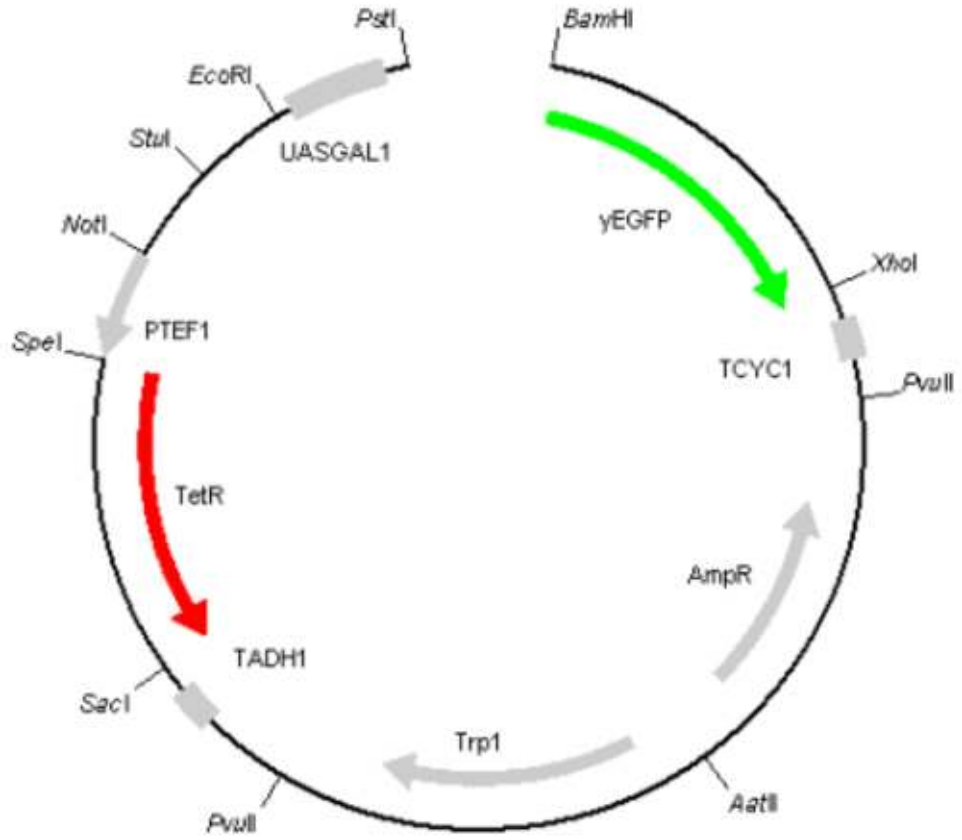
Figures

Figure 1: Schematic for diversity-inherent component synthesis by Klenow pol extension from two annealed long oligos.



Required motif sequences for the function of the component are colored ('site') and unspecified sequences between the motifs are gray ('N').

Figure 2: Example of a characterization vector – for screening and characterization of TetR-repressible yeast promoters.



Features of the vector are separated by unique commonly-used restriction enzyme sites to allow modular construction. Promoter library is placed upstream of the yEGFP reporter coding sequence.

Associated Publications

This protocol is related to the following articles:

- **Diversity-based, model-guided construction of synthetic gene networks with predicted functions**

  See other protocols related to this article

Author information

Affiliations

. **Boston University, Boston**

Tom Ellis

Competing financial interests

The authors declare no competing financial interests.

Readers' Comments

Comments on this thread are vetted after posting.

- ***Protocol Exchange*** ISSN 2043-0116